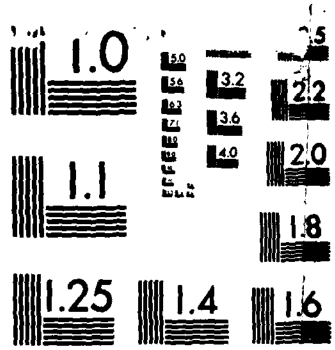
AD-A164 461 THE EFFECT OF LPC (LINEAR PREDICTIVE CODING) PROCESSING 1/1
ON THE RECOGNITION OF UNFAMILIAR SPEAKERS(U) NAVAL
RESEARCH LAB WASHINGTON DC A SCHMIDT-NIELSEN ET AL.
UNCLASSIFIED 20 SEP 85 NRL-8926 F/G 17/2 NL



MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963-A

2

NRL Report 8926

The Effect of LPC Processing on the Recognition of Unfamiliar Speakers

A. SCHMIDT-NIELSEN AND KAREN R. STERN

Information Technology Division
Communication Systems Engineering Branch

September 20, 1985

NAVAL RESEARCH LABORATORY
Washington, D.C.

SECURITY CLASSIFICATION OF THIS PAGE AD-A164 461

REPORT DOCUMENTATION PAGE

1a. REPORT SECURITY CLASSIFICATION UNCLASSIFIED	1b. RESTRICTIVE MARKINGS
2a. SECURITY CLASSIFICATION AUTHORITY 2b. DECLASSIFICATION / DOWNGRADING SCHEDULE	3 DISTRIBUTION / AVAILABILITY OF REPORT Approved for public release; distribution unlimited.
4 PERFORMING ORGANIZATION REPORT NUMBER(S) NRL Report 8926	5. MONITORING ORGANIZATION REPORT NUMBER(S)
6a. NAME OF PERFORMING ORGANIZATION Naval Research Laboratory 6b OFFICE SYMBOL *(If applicable)* Code 7526	7a. NAME OF MONITORING ORGANIZATION
6c. ADDRESS *(City, State, and ZIP Code)* Washington, DC 20375-5000	7b. ADDRESS *(City, State, and ZIP Code)*
8a. NAME OF FUNDING / SPONSORING ORGANIZATION Space and Naval Warfare Sys Cmd 8b. OFFICE SYMBOL *(If applicable)*	9. PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER
8c. ADDRESS *(City, State, and ZIP Code)* Washington, DC 20363-5100	10. SOURCE OF FUNDING NUMBERS PROGRAM ELEMENT NO. 28010N; PROJECT NO.; TASK NO. Z0919 CC; WORK UNIT ACCESSION NO. DN880-072

11. TITLE *(Include Security Classification)*
The Effect of LPC Processing on the Recognition of Unfamiliar Speakers

12. PERSONAL AUTHOR(S)
Schmidt-Nielsen, Astrid and Stern, Karen R.

13a. TYPE OF REPORT	13b. TIME COVERED	14. DATE OF REPORT *(Year, Month, Day)*	15. PAGE COUNT
Interim	FROM ____ TO ____	1985 September 20	15

16. SUPPLEMENTARY NOTATION

17. COSATI CODES		
FIELD	GROUP	SUB-GROUP

18. SUBJECT TERMS *(Continue on reverse if necessary and identify by block number)*
Speaker recognition, Voice distinctiveness
Linear prodictive coding (LPC), Speaker recognition test

19. ABSTRACT *(Continue on reverse if necessary and identify by block number)*

The effect of narrowband digital processing, using a linear predictive coding (LPC) algorithm at 2400 bits/s, on the recognition of previously unfamiliar speakers was investigated. Three sets of five speakers each (two sets of males differing in rated voice distinctiveness and one set of females) were tested for speaker recognition in two separate experiments using a familiarization-test procedure. In the first experiment three groups of listeners each heard a single set of speakers in both voice processing conditions, and in the second two groups of listeners each heard all three sets of speakers in a single voice processing condition. There were significant differences among speaker sets both with and without LPC processing, with the low distinctive males generally more poorly recognized than the other groups. There was also an interaction of speaker set and voice processing condition; the low distinctive males were no less recognizable over LPC than they were unprocessed, and one speaker in particular was actually better recognized over LPC. Although it seems that on the whole LPC processing reduces speaker recognition, the reverse may be the case for some speakers

(Continued)

20 DISTRIBUTION / AVAILABILITY OF ABSTRACT	21. ABSTRACT SECURITY CLASSIFICATION
☒ UNCLASSIFIED/UNLIMITED ☐ SAME AS RPT. ☐ DTIC USERS	UNCLASSIFIED

22a. NAME OF RESPONSIBLE INDIVIDUAL	22b. TELEPHONE *(Include Area Code)*	22c. OFFICE SYMBOL
Astrid Schmidt-Nielsen	(202) 767-2682	Code 7526

DD FORM 1473, 84 MAR 83 APR edition may be used until exhausted. All other editions are obsolete.

19. ABSTRACT (Continued)

in some contexts. This suggests that one should be cautious about comparing speaker recognition for different voice systems on the basis of a single set of speakers. It also presents a serious obstacle to the development of a reliable standardized test of speaker recognizability.

CONTENTS

THE EFFECT OF LPC PROCESSING ON THE RECOGNITION OF UNFAMILIAR SPEAKERS

INTRODUCTION

Narrowband digital voice systems are being increasingly used for secure voice communication applications. A linear predictive coding (LPC) algorithm at 2400 bits/s has been adopted as the government and military standard for this data rate (Federal Standard 1015 or MIL-STD-188-113) as well as by NATO (STANAG 4198). At this low data rate, both the quality and intelligibility of the speech are degraded relative to wideband systems at 64,000 or 32,000 bits/s.

Speaker recognition is one of the aspects that contributes to the quality and acceptability of a voice communication system. It is helpful to be able to recognize the voice of the person you are talking to, whether you are talking over a telephone or using a low data rate (narrowband) digital voice system. Actually, the telephone itself is considerably poorer than the unprocessed comparison speech we used in these experiments. There are also times when it is useful to be able to distinguish the voices of people who were previously unknown to you; for example, in a conference call where one may be conversing with several different speakers at the same time, it is helpful to be able to tell them apart.

It would be highly desirable to have a standardized test procedure (possibly using standard tape recordings with a specified speaker set) that could be used to determine the speaker recognizability for different voice communication systems. Reliable tests for speech intelligibility and quality are available, e.g., diagnostic rhyme test (DRT) [1], modified rhyme test (MRT) [2], and diagnostic acceptability measure (DAM) [3]. Papamichalis and Doddington [4] have proposed a speaker recognizability test in which listeners are asked to identify the speaker of a sentence by comparing it with a series of reference sentences that are continuously available. Their speaker set was composed of five male and five female speakers selected to differ in their confusability with the other speakers in the set. Tests of processed utterances included unprocessed utterances for reference, and both the processed and unprocessed utterances were compared with the unprocessed reference sentences. This form of test can be used to evaluate the fidelity with which a voice processor transmits voice characteristics. Our experience with the telephone suggests that it is possible for people to learn to recognize an individual's processed voice even though it may not be very like the unprocessed voice. A voice system may have high potential speaker recognizability if it transmits information that allows us to discriminate among voices even though it does not reproduce the original voice very well. In this case, a test where the processed voice is the reference would be more appropriate.

In a previous experiment using familiar speakers [5], recognition over the LPC system was approximately 80% of what it was with unprocessed speech from the same speakers. Since most of the listeners were unfamiliar with the LPC system, this result reflects primarily the fidelity of the reproduction. With familiar speakers it is possible to use a reasonably large group of speakers, but this is not feasible with unfamiliar speakers.

It is well recognized that the size and composition of the speaker set have a large effect on recognition performance with previously unknown speakers [6,7]. Practical considerations such as testing time and memory limitations generally make it desirable to limit the speaker set to a relatively small size (Ref. 8 and 9 for a review of speaker recognition test procedures). The continuous comparison

Manuscript approved June 10, 1985.

method [7] used by Papamichalis and Doddington [4] permits a slightly larger set size than the familiarization and test method used by some other investigators [10,11, and 12]. In either case, the small number of speakers means that context effects due to speaker selection will be large and could seriously affect the generality of the test.

Previous investigators have found that there are considerable individual differences in the degree to which different speakers are recognized [7]. The same is also true for intelligibility. Data for intelligibility collected in connection with tests conducted by the Digital Voice Processor Consortium [13] suggest that not only are there individual differences among speakers on intelligibility tests, but it is not necessarily the same speakers who are the most intelligible under different voice processing and noise conditions [14]. In spite of these speaker differences for different voice conditions, the intelligibility test results were consistent in that the voice systems were rank ordered the same for each of the speakers. Hecker and Williams [11] found that for a set of five voice systems, intelligibility and speaker recognition exhibited similar rank order. Unlike intelligibility, speaker recognizability depends not only on the individual voice characteristics, but also on the context of the other speakers in the set and how similar they are to one another. A good test of speaker recognition should be consistent in the same way that an intelligibility test is consistent, namely that voice conditions should be ranked the same across different sets of speakers even though recognition difficulty may vary.

The two experiments described in this report were conducted to investigate the recognizability of unfamiliar speakers talking over a narrowband digital voice communication system, using the DoD standard LPC algorithm, and to compare the effects of different speaker sets in the different test conditions. The consistency of processing effects across different groups of speakers has implications for the generality of any test of speaker recognizability using listener evaluation of small sets of speakers. Rated voice distinctiveness was used to select three groups of five speakers from a set of 24 speakers used in the previous experiment with familiar listeners.

There are several ways in which the LPC system might affect voice characteristics that are related to speaker recognition. The filtering that occurs at frequencies above 3600 Hz removes higher format information that contains important cues to speaker identity. Pitch tracking can be less than perfect and occasional pitch halving or pitch doubling can be confusing. Problems may also occur when there are rapid changes in pitch. Phoneme information tends to be smeared or blurred because of the reduced information rate, as for example, the averaging that occurs over the 22.5-ms frame length. Nonspeech sounds such as coughs, tongue clicks, or lip smacking are not well handled by the algorithm and can be highly distorted. On the other hand, since this is an analysis-synthesis system, prosodic information—rhythm, timing, etc.—remains relatively intact.

EXPERIMENTS

Two experiments were conducted using essentially the same method. The procedure that was selected was a familiarization phase followed by a test phase rather than the continuous comparison procedure. In the first experiment, the listeners heard both the processed and unprocessed version for the same set of speakers in counterbalanced order. Since there were large differences in the listeners' ability to recognize speakers, this design reduced chance effects of listener variability on the differences due to processing, but a particularly good or poor listener might have an effect on the speaker group differences. This design could also be susceptible to differential practice effects since the same speakers were heard twice, once in each processing condition. In the second experiment the same listeners were tested on all three speaker sets but heard only a single version, LPC processed or unprocessed. This design complements that of the first experiment in that the effects of individual differences on speaker sets were controlled, and practice effects were minimized since successive tests involved different speakers.

General Method

Speakers and Speech Materials

Three sets of five speakers were selected from a group of 24 speakers used in a previous experiment [5]. There were two sets of male speakers; the first group consisted of speakers who had been rated as having more distinctive or characteristic voices and the second group was rated as having less distinctive voices (these will be referred to as the high males and the low males). For the voices in this experiment, we had two independent sets of distinctiveness ratings—one by 24 people who knew the speakers and one by 54 listeners unfamiliar with the speakers, none of whom were listeners in the present experiments. Both groups used a 7-point scale to answer the question *How distinctive or characteristic is this person's voice*? The familiar ratings were done from memory, and the unfamiliar raters heard tape recorded voice samples. The male voices were assigned to two groups according to the average of the two sets of distinctiveness ratings. The third group consisted of five female voices varying in distinctiveness (there were not enough females for two groups). Speech samples from the speakers talking in a conversational manner were taken from the materials used in the previous experiment and consisted of excerpts from recordings of pairs of speakers playing a game of battleship [15]. The battleship game provided the opportunity for two speakers seated in separate sound booths to communicate with one another in a natural manner, and at the same time ensured a reasonably consistent vocabulary for the different speakers since the vocabulary needed to play the game is quite limited—naming squares in the playing grid; for example, *My shot is bravo two,* or giving responses; for example, *That's a miss.* The speakers were recorded playing together in pairs. Games were recorded in two separate sessions, one over an unprocessed voice channel and the other with two players talking over the LPC voice processor. Thus the speakers could talk the way they normally would for each type of voice channel. This meant that it was possible to compensate for the poorer quality of the LPC system by talking more slowly and carefully, as one would do when using the system in real-life situations. The battleship games were spliced apart to obtain a number of excerpted phrases for each speaker. There were no significant differences among speakers in the average duration of the selected phrases, although the LPC phrases (mean, 2.2 s) were slightly longer than the unprocessed phrases (mean, 2.0 s), owing to the tendency to speak more slowly and carefully when talking over the LPC processor. Each speaker was also recorded reading two familiarization paragraphs, one for the unprocessed condition and one for the LPC condition. Each paragraph lasted about 30 s, and both contained approximately the same number of words. The fact that the familiarization paragraphs were read whereas the test materials were conversational may have made the identifications more difficult, but it was not considered feasible to try to collect 30 s of highly comparable spontaneous speech from each of 15 different speakers. Instead all speakers read the same paragraphs to ensure that the familiarization materials were comparable.

Procedure

The experiment consisted of a familiarization phase in which the speakers' voices were introduced followed by a test phase during which the listeners tried to identify the conversational phrases spoken by the different speakers. In the familiarization phase, each speaker introduced himself or herself giving a fictitious name starting with one of the letters from A to E, by saying, *Hello, my name is* ____ and then reading the familiarization paragraph. The paragraph for the unprocessed condition was about quicksand and was presented unprocessed; the one for the LPC condition was about a Chinese restaurant and was LPC processed. To minimize confusion for the listeners, the familiarization paragraphs were always presented by speakers in order from A to E. The listeners were given typed copies of the test so that they could concentrate on the voice rather than the content. The five paragraphs were followed by a practice test of five phrases, one for each speaker, given in random order with feedback at the end. At this point the difficulty of the task became apparent to the subjects, and familiarization was repeated. The test phase consisted of 25 conversational excerpts, five for each speaker, presented in

pseudorandom, counterbalanced order. Each excerpt was preceded by a 1000 Hz tone and was followed by 4 s of silence during which the subjects wrote the letter corresponding to the speaker's name on a numbered answer sheet and checked a confidence rating of *very sure, fairly sure,* or *guessing.* The subjects were instructed not to leave any blanks and to guess if they had to. The subjects were tested in groups of from 1 to 5 and heard the test tapes in a quiet room over high quality headphones.

Experiment 1

Method

Volunteers unfamiliar with any of the speakers were recruited through the University of Maryland Psychology Department. There were 72 listeners, 24 for each of the three groups of speakers. All subjects heard both an unprocessed and an LPC processed tape of the same speakers. One-half the subjects were familiarized and tested on the unprocessed condition first, and for the other half the order was reversed.

Results

Figure 1 shows the percent of correct responses for each of the three groups. The dotted line indicates chance performance. Analysis of variance [16] showed a significant effect of speaker sets, $F(2,66) = 3.83$, $p < 0.05$. Recognition of the high males and the females was considerably better than the low males. Speaker recognition over LPC was significantly poorer than with unprocessed speech, $F(1,66) = 37.80$, $p < 0.001$. The Tukey test for differences between means [16] showed that this difference was significant for the high males and the females. The low males were actually recognized slightly better over LPC than they were unprocessed, but this difference was not statistically significant, although there was a significant speaker group by processing condition interaction, $F(2,66) = 24.32$, $p < 0.001$. There was also a significant learning effect over trials, $F(1,66) = 20.07$, $p < 0.001$, although there seemed to be less improvement if the LPC condition preceded the unprocessed than the other way around.

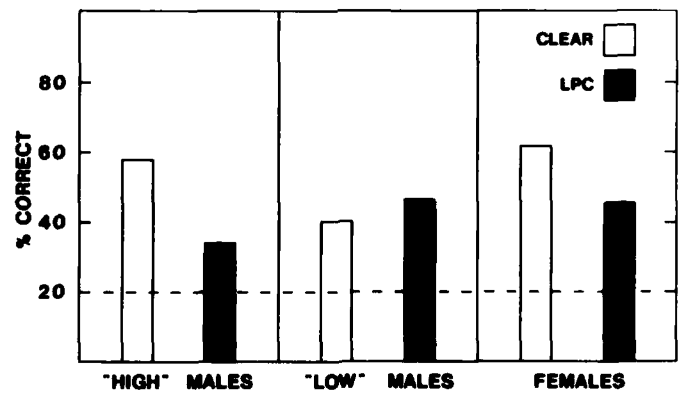


Fig. 1 — Speaker group scores for unprocessed and LPC processed speech for Experiment 1

Figure 2 shows the individual results for the speakers in each set. The speakers are shown from left to right by the code letters that were the initials of the made-up names. For each speaker set the results are consistent with the results for the group as a whole. All five of the high males showed a large loss in recognizability with LPC processing. The female speakers also showed a loss in recognition for all five speakers- more for some than for others. The five low males had an entirely different pattern. No speaker showed any significant drop in recognition due to LPC, and two seem to have improved — one speaker, Bob, accounts for most of the real gain that was seen for this group. It is not

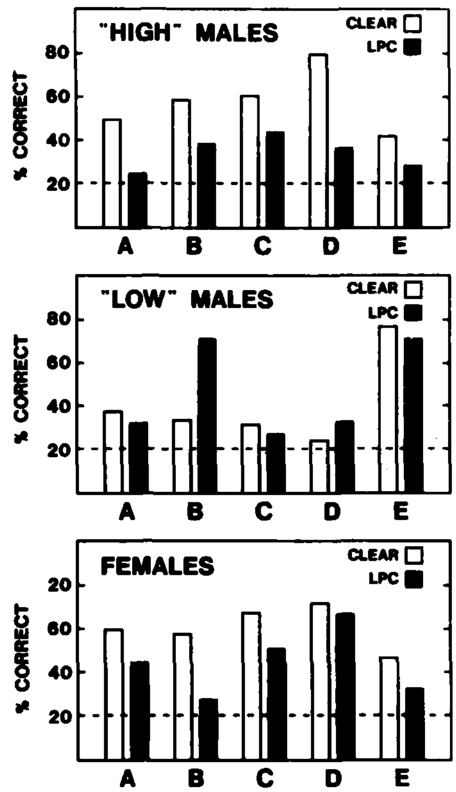


Fig. 2 — Individual speaker scores for unprocessed and LPC processed speech for Experiment 1

clear exactly how this effect is related to voice distinctiveness since within speaker groups, there was no consistent relationship between the rated distinctiveness of a particular voice and the recognition of that voice. In fact, one of the two best recognized female voices, Carol, was also the one rated the least distinctive.

Experiment 2

Method

The subjects were 19 psychology students recruited at the University of Maryland during the summer session. Each subject was tested with all three speaker sets but heard only one version, LPC or unprocessed. There were 9 listeners for the unprocessed and 10 for the LPC version. Because of the difficulty of obtaining subjects during the summer, the order in which the speaker sets were presented to the listeners was balanced for the unprocessed condition, but it was not fully balanced for the LPC condition. Fortunately, post hoc tests showed no significant effect of test order. In the first experiment, speaker E for the high males (Edward) was relatively poorly recognized while speaker E for the low males (Eric) was very well recognized. For the second experiment these two speakers were exchanged so that Eric was in the high male group and Edward was placed in the low male group. This manipulation should have the effect of increasing the difference between the two groups.

Results

Figures 3 and 4 show the comparison between the three sets of speakers and the individual speaker scores. The scores were slightly lower than in the previous experiment because the same

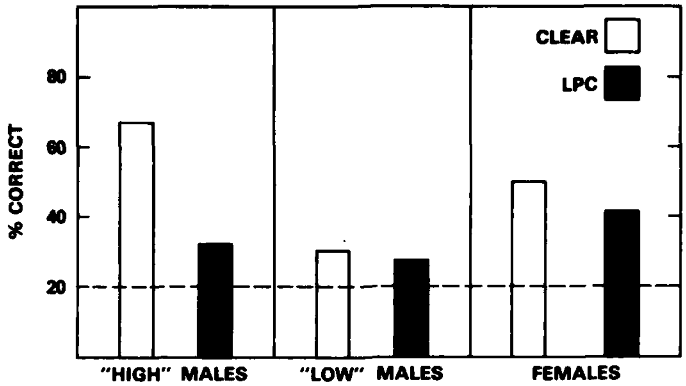


Fig. 3 — Speaker group scores for unprocessed and LPC processed speech for Experiment 2

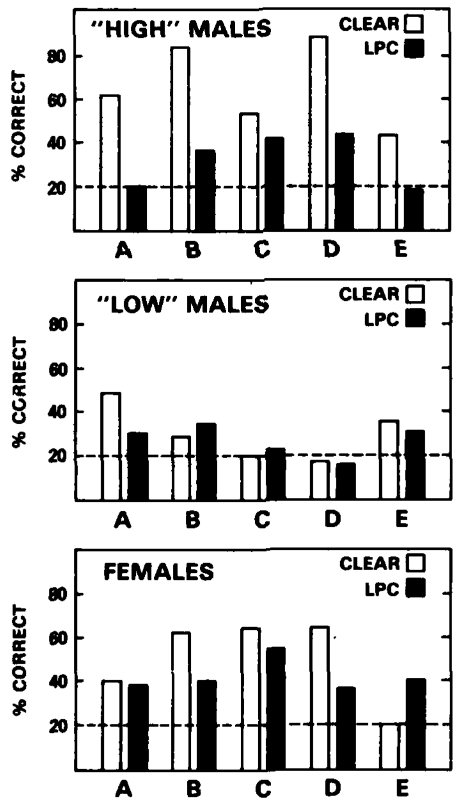


Fig. 4 — Individual speaker scores for unprocessed and LPC processed speech for Experiment 2

speakers were only heard once. Analysis of variance showed that recognition of the high males and the females was again considerably better than the low males, $F(2,34) = 7.94$, $p < 0.01$. Exchanging Eric and Edward had the expected effect of increasing the difference between the high and low males, and the high males were now recognized better than the females. Speaker recognition over LPC was significantly poorer than with unprocessed speech, $F(1,38) = 11.98$, $p < 0.01$, and there was a significant speaker group by processing condition interaction, $F(2,34) = 4.44$, $p < 0.05$. The low males in this experiment were recognized slightly but insignificantly worse over LPC than unprocessed, and there was only a small improvement for Bob over LPC. This change can probably be attributed to the fact that Bob was frequently confused with Edward in the second experiment whereas there were no confusions of Bob with Eric in the first experiment. These changes in the pattern of results due to exchanging one pair of speakers again emphasize the extreme dependence of recognition scores on the composition of the speaker set when small groups of unfamiliar speakers are used.*

DISCUSSION AND CONCLUSIONS

In both experiments the composition of the speaker set affected the overall recognition rate, and there was also an interaction with processing condition. The two sets of male voices were originally grouped by rated voice distinctiveness and not by any direct measure of the similarity of the voices in each group. It could be that the more distinctive voices were easier to tell apart because each voice was unusual in its own way, whereas the less distinctive voices were all more ordinary.

It is not surprising that LPC processing and the accompanying loss of information should make the voices less distinct from one another, and this is what happened for the high males and the females, but not for the low males. In the earlier experiment using listeners who were familiar with the speakers [5], the recognition of the individual speaker was uncorrelated with distinctiveness ratings (either by familiar or unfamiliar raters). It is more likely that voice distinctiveness should be a factor in the recognition of unfamiliar speakers than of known speakers. The results of the present experiments, however, indicate that although grouping the speakers by rated distinctiveness had a significant effect on recognition of the group as a whole, the recognition of individual speakers was again uncorrelated with rated distinctiveness.

Voice distinctiveness does seem to have an effect on speaker recognition, but the nature of the relationship is unclear. One problem may be in the inconsistency of the rating process as there was little agreement among raters for most of the speakers. Different listeners may have different concepts in mind as they perform the rating task. A voice can be distinctive in many ways. For example, it may be distinctive in a particular context (e.g., the only female in a group of males), but some voices also seem to be inherently more distinctive than others (e.g,, a voice one feels one would recognize anywhere). Further research is needed on the relationship between rated voice characteristics and speaker recognition as this is a problem that has proved difficult to resolve. It may be that the use of more specific questions would provide better answers.

The female speaker set was more heterogeneous with respect to distinctiveness, with one very high rating and one extremely low rating, than were the two sets of male speakers. Recognition of this

*The recognition results for the low males in the first experiment and for Bob in particular do not seem to have been simply chance fluctuations since the pilot study for this experiment showed a similar pattern of results, although the scores were slightly lower because the familiarization paragraphs were only heard once. The scores for the pilot study were:

- high males — unprocessed, 47%, LPC 37%;
- females — unprocessed, 42%, LPC 31% ;
- low males — unprocessed, 30%, LPC 39%, and
- Bob — unprocessed, 35%, LPC 60%.

(One of the authors met Bob at a Halloween costume party and completely failed to recognize him from his voice in spite of knowing him well from work. He seems to have a very anonymous sounding voice that becomes more distinct from other voices when it is heard over LPC. The voice did not sound odd or distorted in the LPC condition.)

mixed group was more similar to the high males than the low males in both experiments, and it is possible that a mixed group would be more representative of overall performance with a larger population. Still, the fact that there was no recognition loss for the low males argues for extreme caution in drawing general conclusions on the basis of a small group of speakers. The females on the average were rated lower in distinctiveness than the two groups of male speakers. Since this could reflect a bias in the way men and women are perceived, it is perhaps best to avoid making direct comparisons between the different sex groups regarding the effects of distinctiveness.

Recognition in the second experiment was somewhat lower than in the first, where the same speakers were heard in both conditions. Figure 5 illustrates the effect of trials. It can be seen that most of the improvement in the first experiment occurred when the unprocessed condition preceded the LPC condition rather than the other way around. This suggests that in addition to experience with the LPC processed voice, knowing a speaker's unprocessed voice is helpful in learning to recognize that person's LPC voice. Recognition of the low males in the second experiment, after exchanging voices of Edward and Eric, was quite poor, only slightly better than guessing, in both the unprocessed and the LPC condition. This suggests the possibility of a floor effect, which could be a reason that the scores did not drop in the LPC condition; however, a binomial test showed that the scores in both conditions were significantly above chance.

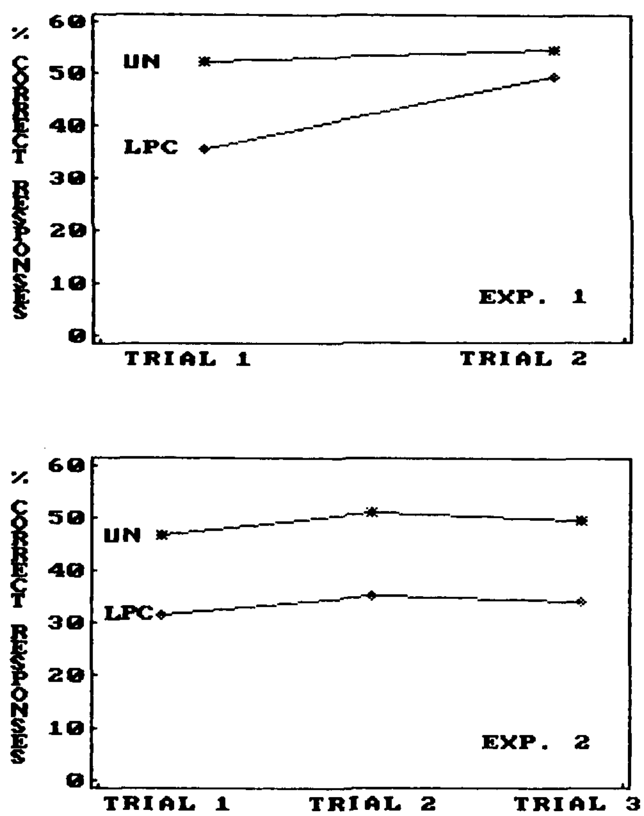


Fig. 5 — The effects of practice in the two experiments. Speakers were the same and processors were different on separate trials in Experiment 1, and processors were the same and speakers were different on separate trials in Experiment 2.

The overall recognition rate using the familiarization-test procedure was considerably lower than that for the familiar listeners used in the previous experiment. This is partly due to the memory problems inherent in learning a new set of voices. The difference between the training materials, which were read, and the test materials, which were conversational, may have made the task especially difficult. However, Legge et al. [12] using an old-new paired comparison task also obtained low recognition rates even though both familiarization and test materials were read. These investigators comment that recognizing a person by voice alone is a particularly difficult task.

There are a number of problems to be solved in developing a standardized test of speaker recognition. Such a test must for practical reasons rely on the use of previously unknown speakers. This means that realistically the size of the speaker set will be relatively small because of the constraints of such factors as memory load, training, and testing time. The present results suggest that with small sets of speakers, the composition of the speaker set is extremely important. Not only did the scores for individual speakers change depending on the context of the group, but the effect of LPC processing was different for different speaker sets. Considerable research is needed to determine whether it is possible to select a set of speakers (or possibly several sets) that will give results that are reasonably representative of the performance that can be expected with a larger population and that are consistent for a variety of different voice processing conditions. It may be that a continuous recognition task is not as susceptible to speaker variation, but the results of Stevens et al. [17] suggest that this is not the case. Perhaps other methods of evaluating speaker recognition should be considered, for example, voice rating scales [18,19]. However ratings have so far not been shown to discriminate among speakers as well as direct listening methods [20]

It seems reasonable to conclude that on the whole the effect of LPC processing is to reduce speaker recognizability but that this is not necessarily the case for all speakers and can be highly context dependent. The two groups that were well recognized in the unprocessed condition showed losses in recognition over the LPC system that were similar to the loss for the familiar speakers in the previous experiment, whereas the group that was poorly recognized on the unprocessed condition showed no further loss under LPC processing. This suggests that while there is clearly a loss in the fidelity with which the voice is transmitted, there is still some potential for discriminating among voices heard over the LPC system. There are large and real differences among speakers in recognition over the LPC system. The potential recognition of some may be quite high once their "LPC voice" is learned whereas others lose some of their distinctiveness and are harder to recognize.

REFERENCES

1. W.D. Voiers, "Diagnostic Evaluation of Speech Intelligibility," *Speech Intelligibility and Recognition*, M.E. Hawley, ed. (Dowden, Hutchinson, and Ross, Stroudsburg, PA, 1977).

2. A.S. House, C.E. Williams, M.H.L. Hecker, and K.D. Kryter, "Articulation-Testing Methods: Consonantal Differentiation with a Closed-Response Set," *J. Acoust. Soc. Am.* 37, 158-166 (1965).

3. W.D. Voiers, "Diagnostic Acceptability Measure for Speech Communication Systems, ICASSP 77. 1977 IEEE International Conference on Acoustics, Speech, and Signal Processing, New York, 1977.

4. P.E. Papamichalis and G.R. Doddington, "A Speaker Recognizability Test," ICASSP 84. 1984 IEEE International Conference on Acoustics, Speech, and Signial Processing, San Diego, CA, 1984.

5. A. Schmidt-Nielsen and K.R. Stern, "Identification of Known Voice as a Function of Familiarity and Narrowband Processing," *J. Acoust. Soc. Am.* 77 (2), 658-663 (1985).

6. I. Pollack, J.M. Pickett, and W.H. Sumby, "On the Identification of Speakers by Voice," *J. Acoust. Soc. Am.* **26**, 403-406 (1954).

7. K.N. Stevens, C.E. Williams, J.R. Carbonell, and B. Woods, "Speaker Authentication and Identification: A Comparison of Spectrographic and Auditory Presentations of Speech Material," *J. Acoust. Soc. Am.* **44**, 1596-1607 (1968).

8. P.D. Bricker and S. Pruzansky, "Speaker Recognition," in *Contemporary Issues in Experimental Phonetics*, N.J. Lass, ed. (Springfield, IL, Thomas, 1975), pp. 295-326.

9. M.H.L. Hecker, "Speaker Recognition: an Interpretive Survey of the Literature," *ASHA Monogr. 16* (American Speech and Hearing Association, Washington, DC, 1971).

10. P.D. Bricker and S. Pruzansky, "Effects of Stimulus Content and Duration on Talker Identification," *J. Acoust. Soc. Am.* **40**, 1441-1449 (1966).

11. M.H.L. Hecker and C.E. Williams, "On the Interrelation Among Speech Quality, Intelligibility, and Speaker Identifiability," Proceedings of the Fourth International Acoustics Congress, Liege, Belgium, 1965, pp. 1-4.

12. G.E. Legge, C. Grosmann, and C.M. Peiper, "Learning Unfamiliar Voices," *J. Exp. Psychol.: Learn. Mem. Cogn.* **10**, 298-303 (1984).

13. G.F. Sandy, "Digital Voice Processor Consortium Final Report," MTR-84W00053, MITRE Corp, McLean, VA, March 1984.

14. A. Schmidt-Nielsen, "Tests of Usability and Intelligibility of LPC in Different Environments," NRL Report 8851, 1984.

15. A. Schmidt-Nielsen and S.S. Everett, "A Conversational Test for Comparing Voice Systems Using Working Two-way Communication Links," *IEEE Trans. Acoust. Speech Sig. Proc.* **30**, 853-863 (1982).

16. B.J. Winer, *Statistical Principles in Experimental Design* (McGraw-Hill, New York, 1971), 2nd ed.

17. K.N. Stevens, H.L. Heckler, and K.D. Kryter, Rome Air Force Base Contact No. AF 30 (602-2235) (1962), RADC-TDR-62-171.

18. W.D. Voiers, "Perceptual Bases of Speaker Identity," *J. Acoust. Soc. Am.* **36**, 1065-1073 (1964).

19. W.D. Voiers, "Toward the Development of Practical Methods of Evaluating Speaker Recognizability," ICASSP 79. 1979 International Conference on Acoustics, Speech, and Signal Processing, Washington, DC, April 1979. Record. (IEEE, New York), pp. 793-796.

20. F.R. Clarke and R.W. Becker, "Comparison of Techniques for Discriminating Among Talkers," *J. Speech Hear. Res.* **12**, 747-761 (1969).

END

FILMED

3-86

DTIC